Stat 543 4-29-05

Course Eval . . . 1/A Bad
5/E Good

Recall

Lemma

$$Q_n(\Delta) = \ell_n(\hat{\theta}_n) - \ell_n\left(\hat{\theta}_n + \frac{\Delta}{\sqrt{n}}\right) \xrightarrow{P_\theta} \frac{1}{2}\Delta^2 I_1(\theta)$$

(The loglikelihood is locally quadratic near its maximum)

Pf: Do a Taylor expansion of $l_n(\cdot)$ around $\hat{\theta}_n$

$$l_n\left(\hat{\theta}_n + \frac{\Delta}{\sqrt{n}}\right) - l_n(\hat{\theta}_n) = \cancel{\frac{\Delta}{\sqrt{n}} l_n'(\hat{\theta}_n)} \to 0$$

$$+ \left(\frac{\Delta}{\sqrt{n}}\right)^2 \frac{1}{2} l_n''(\hat{\theta}_n)$$

$$+ \frac{1}{6}\cancel{\left(\frac{\Delta}{\sqrt{n}}\right)^3 l_n'''(\theta_1)} \to$$ goes to 0 in P_θ probability

for θ_1 some value between $\hat{\theta}_n$ and $\hat{\theta}_n + \frac{\Delta}{\sqrt{n}}$

$$\frac{1}{2}\Delta^2 \, \frac{1}{n} l_n''(\hat{\theta}_n) \xrightarrow{P_\theta} I_1(\theta)$$

Corollary Supposing a prior dsn G has a pdf $g(\cdot)$ on $\Theta \subset \mathbb{R}^1$, $g(\theta_0) > 0$ and $g(\cdot)$ cont^s at θ_0 — If $\hat{\theta}_n$ is an MLE of θ and is consistent at θ_0 the posterior density $g(\theta | X)$ has the property that

$$R(\Delta) = \ln \frac{g(\hat{\theta}_n | X)}{g(\hat{\theta}_n + \frac{\Delta}{\sqrt{n}} | X)} \xrightarrow{P_{\theta_0}} \frac{1}{2} \Delta^2 I_1(\theta_0)$$

Pf:

$$R(\Delta) = \ln \frac{g(\hat{\theta}_n) L_n(\hat{\theta}_n)}{g(\hat{\theta}_n + \frac{\Delta}{\sqrt{n}}) L_n(\hat{\theta}_n + \frac{\Delta}{\sqrt{n}})}$$

$$= \ln g(\hat{\theta}_n) - \ln g(\hat{\theta}_n + \frac{\Delta}{\sqrt{n}}) \xrightarrow{P_{\theta_0}} 0$$

$$+ l_n(\hat{\theta}_n) - l_n(\hat{\theta}_n + \frac{\Delta}{\sqrt{n}}) \longleftarrow Q_n(\Delta)$$

∎

It says that near the MLE the shape of the log posterior density is the same as the shape of the log likelihood, i.e. quadratic (so shape of posterior is approximately normal !)

Why?

Suppose Y is a cont's r.v. with pdf f and

$$\ln \frac{f(0)}{f(y)} = \frac{c}{2} y^2$$

What is the dsn of Y?

$$\frac{f(0)}{f(y)} = \exp \frac{c}{2} y^2$$

$$f(y) = f(0) \exp - \frac{c}{2} y^2$$

Y is normal with mean 0 and variance $\frac{1}{c}$

So posteriors often look normal ... This follows from the fact that log-likelihood look quadratic for "large n"

Very short version of LRT, Wald, Score tests –

$$\underset{k\times 1}{\theta} = \begin{pmatrix} \underset{l\times 1}{\theta_1} \\ \underset{(k-l)\times 1}{\theta_2} \end{pmatrix} \quad \text{for Testing} \quad H_0 : \theta_1 = \theta_{01}$$

<u>LRT</u> Test Statistic

$$2\left(\underbrace{\ell_n(\hat{\theta}_n)}_{\sup_{\theta} \ell_n(\theta)} - \underbrace{\ell_n^*(\theta_{01})}_{\sup_{\theta_2} \ell_n(\theta_{01},\theta_2)}\right)$$

χ^2_ℓ limit under H_0

must find an MLE

must find a "restricted" MLE

<u>Wald</u> Test Statistic

$$\underbrace{\left(\hat{\theta}_{n1} - \theta_{01}\right)'\left(\text{est'd covariance matrix for } \hat{\theta}_{n1}\right)^{-1}\left(\hat{\theta}_{n1} - \theta_{01}\right)}$$

χ^2_ℓ limit under H_0

This requires computation of $\hat{\theta}_n$ (the MLE)

In fact, at end of the world these tests are equivalent

(Rao) Score Test motivation here is that if H_0 is true and $\tilde{\theta}_{n2}(\theta_{01})$ is the maximizer of $l_n(\theta_{01}, \cdot)$ then the score function at $(\theta_{01}, \tilde{\theta}_{n2}(\theta_{01}))$ should be close to 0 (score is 0 at MLE) and score test statistic is an appropriate quadratic form in

$$\left(\frac{\partial}{\partial \theta_i} l_n(\theta) \Big|_{\theta = \begin{pmatrix} \theta_{01} \\ \tilde{\theta}_{n2}(\theta_{01}) \end{pmatrix}} \right) \leftarrow k \times 1$$

To apply a score test you must do a restricted maximization of likelihood — an appropriate quadratic form is also χ^2_ℓ in the limit under H_0 —

At the end of the world these tests are equivalent